Tailweight with respect to the mode for unimodal distributions

J. Averous, A.-L. Fougères*, M. Meste

Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118 Rte de Narbonne, 31062 Toulouse cedex, France

Abstract

Location, spread, skewness and tailweight are studied for unimodal distributions by means of mode-based concepts. The Lévy concentration function and notions related to it are playing an important part.

1. Introduction

Unimodal distributions form a remarkable subset of probability distributions, which presents nice properties. Several characterizations of unimodal distributions can be found in the literature: concentration functions (Bertin et al., 1981), characteristic functions (e.g. Dharmadhikari and Joag-Dev, 1988, p. 7). In Bertin et al.'s approach to unimodality, the notions of concentration function Q_F and pointer A_F of a distribution function F are essential.

Location, scale, skewness and tailweight are important concepts for the description of a probability distribution. The study of tails for skewed distributions presents a particular aspect for unimodal distributions: when we consider the graph of the probability density function, the mode seems to be an appealing centre.

An analogous approach to the quantile-based description of F is envisaged; it takes the mode as reference and makes use of A_F and Q_F as location and dispersion functional parameters. The mode, as location parameter, does not preserve the stochastic ordering. However, its interpretation makes it more rational in certain circumstances than mean or median. So, a description with respect to the mode of a distribution would be welcomed.

Basic definitions and notations are given in Section 2. Section 3 then deals with the concept of location with respect to the mode by making use of the modal central interval and its midpoint. A way to approach distributions in the mode-sense, using already known notions in the median-sense, is suggested. Section 4 concerns tailweight orderings and skewness. A location- and scale-free tailweight ordering is introduced to

---

* Corresponding author.

compare tail with respect to the mode, without assuming symmetry. Finally, mode-based concepts of qualitative aspect of skewness are considered.

2. Preliminaries

Let F be a (right continuous) distribution function on $\mathbb{R}$. F is said to be *unimodal* whenever there is an a, called a *mode* of F, such that F is convex on $(-\infty, a)$ and concave on (a, ∞). Further let $x' = \inf\{x: F(x) > 0\}$ and $x'' = \sup\{x: F(x) < 1\}$; F is said to be *strictly unimodal* (Bertin et al., 1981) if it is unimodal with mode a and if it is strictly convex on (x', a) and strictly concave on (a, x''); here $S_F = [x', x'']$ is the support of F.

Let us recall the definitions and properties of the Lévy concentration function and pointer: let F be a continuous distribution function; then the *Lévy concentration function* of F is defined by:

$$Q_F(\lambda) = \begin{cases} \sup_{x \in \mathbb{R}} (F(x + \lambda) - F(x)) & \text{for } \lambda \geqslant 0, \\ 0 & \text{for } \lambda < 0. \end{cases}$$

Next the *pointer A_F* of F is the map from $\mathbb{R}_+$ to the set of all subsets of $\mathbb{R}$, defined by:

$$A_F(\lambda) = \{x \in \mathbb{R}: Q_F(\lambda) = F(x + \lambda) - F(x)\}.$$

Let $\mathcal{U}$ be the set of all unimodal continuous distribution functions; then $F \in \mathcal{U}$ if and only if $Q_F \in \mathcal{U}$ and $A_F(\lambda)$ is an interval for each $\lambda \in \mathbb{R}_+$ (Bertin et al., 1981). If F is strictly unimodal, then $A_F(\lambda)$ is a singleton, say $\{x_\lambda\}$.

Let $F \in \mathcal{U}$; if several modes exist, they form an interval of $\mathbb{R}$ and let M_F be the midpoint of this interval. If δ denotes the length of $A_F(\lambda)$, the *modal central interval* of length $\lambda \in \mathbb{R}_+$ of F is defined by:

$$I^0(\lambda, F) = [\inf A_F(\lambda) + \delta/2, \ \inf A_F(\lambda) + \delta/2 + \lambda].$$

For F strictly unimodal, we have $I^0(\lambda, F) = [x_\lambda, x_\lambda + \lambda]$.

Let us now introduce a continuous distribution function $\widetilde{F}$, which helps to transfer mode-based notions in quantile-based ones; for each unimodal distribution function F with continuous and bounded probability density function f, we define:

$$\widetilde{F}(x) = \begin{cases} \dfrac{f(x)}{2f(M_F)} & \text{for } x \leqslant M_F, \\ 1 - \dfrac{f(x)}{2f(M_F)} & \text{for } x > M_F. \end{cases}$$

Clearly, if we note $I^1(\lambda, F)$ the interquantile interval $[F^{-1}(\alpha), F^{-1}(1 - \alpha)]$ with length λ, we have $I^0(\lambda, F) = I^1(\lambda, \widetilde{F})$, and particularly, for $\lambda = 0$, $\text{med}(\widetilde{F}) = M_F$.

3. Mode-based concept of location

Let $c^0(\lambda, F)$ be the midpoint of $I^0(\lambda, F)$; eventually we shall replace F by X, X being a random variable with distribution function F. The map $F \mapsto c^0(\cdot, F)$, viewed as a generalization of a location parameter, has the following equivariance property:

Proposition 3.1.

$$c^0(\cdot, aX + b) = ac^0(\cdot/|a|, X) + b \quad \textit{for each } a \neq 0.$$

Proof. It follows from the equivariance property of the pointer $A_F(\cdot)$. $\square$

Generally, location parameters are obtained by optimisation criteria. For example, the median appears as minimizing $\mathbb{E}(|X - c|)$, $c \in \mathbb{R}$, where X is an integrable random variable. The modal central interval, viewed as a functional location parameter, may be obtained in the same way: for each $\lambda > 0$,

$$\mathbb{E}(d_0(X, I^0(\lambda, X))) = \min\{\mathbb{E}(d_0(X, I)) : I \in \mathscr{I}_\lambda\},$$

where $d_0(x, y) = 0$ if $x = y$, and 1 otherwise, and $\mathscr{I}_\lambda$ denotes the set of all intervals of length λ. Indeed, we have $\mathbb{E}(d_0(X, I)) = 1 - P(I)$, and so $\min\{\mathbb{E}(d_0(X, I)) : I \in \mathscr{I}_\lambda\} = 1 - \max\{P([x, x + \lambda]) : x \in \mathbb{R}\}$. Then definitions of $A_F(\lambda)$ and $I^0(\lambda, F)$, and continuity of F imply the result. In other words, this is just another way to say that $I^0(\lambda, F)$ is the interval $(a, a + \lambda)$ for which the concentration function is reached, i.e. $Q_F(\lambda) = F(a + \lambda) - F(a)$. If several solutions exist, i.e. $A_F(\lambda)$ is not a singleton, then the pointer forms an interval (Bertin et al., 1981), and we consider its midpoint, like in the definition of $I^0(\lambda, F)$.

The following result concerns the map $(\lambda, F) \mapsto c^0(\lambda, F)$:

Proposition 3.2. (i) *Let* $F_n, F \in \mathscr{U}$, $n \in \mathbb{N}$, *and let* F *be strictly unimodal. Assume that* $F_n \xrightarrow{w} F$; *then* $c^0(\lambda, F_n) \to c^0(\lambda, F)$ *for each* $\lambda \in \mathbb{R}_+$.

(ii) *For each* $F \in \mathscr{U}$, $c^0(\cdot, F)$ *is continuous on* $\mathbb{R}_+$.

Proof. (i) Let us consider a fixed $\lambda \in \mathbb{R}_+$; since F is a strictly unimodal distribution function, $A_F(\lambda) = \{x\}$. Let us show that each sequence $(x_n)_{n \in \mathbb{N}}$, $x_n \in A_{F_n}(\lambda)$, $n \in \mathbb{N}$, converges to x. Strict unimodality of F implies that each $(x_n)_{n \in \mathbb{N}}$ is convergent, note y the limit. Then $y = x$; indeed, $F_n \xrightarrow{w} F$ implies $Q_{F_n} \xrightarrow{w} Q_F$ and then continuity of F_n and F leads to $y \in A_F(\lambda) = \{x\}$.

(ii) is an immediate consequence of continuity properties of $\lambda \mapsto \inf A_F(\lambda)$ and $\lambda \mapsto \sup A_F(\lambda)$ (Bertin et al., 1981). $\square$

In the following, we restrict ourselves to unimodal distribution functions with continuous and bounded probability density function. Now, let us define the notion of tailweight with respect to the mode. The median central interval $I^1(\lambda, F)$ has the property of leaving right and left tails of same weight, which can be controlled by the maps $x \mapsto F(x)$ and $x \mapsto 1 - F(x)$, respectively. Since the probability density function f has the same value at the endpoints of the modal central interval, we may say that f controls the tailweight from the mode-based point of view. The relationship between $\widetilde{F}$ and F leads to the description of the tailweight by means of mode-based notions. For example, let us consider the couple of functional parameters (m_F, λ_F), defined, for each $u \in (0, \frac{1}{2})$, by:

$$m_F(u) = \tfrac{1}{2}(F^{-1}(u) + F^{-1}(1 - u)), \qquad \lambda_F(u) = |F^{-1}(1 - u) - F^{-1}(u)|,$$

where $F^{-1}(u) = \tfrac{1}{2}(\inf\{x : F(x) \geqslant u\} + \sup\{x : F(x) \leqslant u\})$. m_F is a location parameter and λ_F a dispersion parameter (Bickel and Lehmann, 1979) for a description of F in terms of tailweight. Another possibility is to use the couple $(c^1(\cdot, F), \alpha^1(\cdot, F))$, where $c^1(\lambda, F)$ is the centre of $I^1(\lambda, F)$ and $\alpha^1(\lambda, F) = 1 - \text{Prob}(I^1(\lambda, F))$, for a description in terms of the length of the central interval (Averous and Meste, 1990).

By analogy, we may define two couples of parameters of location and dispersion with respect to the mode; the first one is (m_F^0, λ_F^0), where $m_F^0 = m_{\widetilde{F}}$ and $\lambda_F^0 = \lambda_{\widetilde{F}}$. Note that we have:

$$\lambda_F^0(u) = \bar{f}_+^{-1}(u) - \bar{f}_-^{-1}(u) \quad \text{and} \quad m_F^0(u) = \tfrac{1}{2}(\bar{f}_+^{-1}(u) + \bar{f}_-^{-1}(u)),$$

where $\bar{f}(x) = f(x)/2f(M_F)$, $\bar{f}_-^{-1}(u)$ is the centre of $\{x \in (-\infty, M_F] : \bar{f}(x) = u\}$ and $\bar{f}_+^{-1}(u)$ is the centre of $\{x \in [M_F, +\infty) : \bar{f}(x) = u\}$.

Let now $\alpha^0(\cdot, F)$ be the map from $\mathbb{R}_+$ to $(0,1]$ defined by $\alpha^0(\lambda, F) = \alpha^1(\lambda, \widetilde{F})$. Then the couple $(c^0(\cdot, F),$ $\alpha^0(\cdot, F))$ enables us to describe location and dispersion in terms of length of the modal central interval. We also obtain the following relationships:

$$m_F^0(u) = c_F^0 \circ \lambda_F^0(u), \quad u \in (0, \tfrac{1}{2}],$$

$$\alpha^0(\lambda, F) = 2(\lambda_F^0)^{-1}(\lambda), \quad \lambda \in \mathbb{R}_+.$$

Unlike the median-based approach, where $\alpha^1(\cdot, F)$ (or its inverse) is often used for dispersion, the concentration function (more exactly $\rho : \lambda \mapsto 1 - Q_F(\lambda)$) seems to be in the mode-sense more natural for the description of F. We have:

$$\alpha^0(\cdot, F) = - \frac{\rho'}{f(M_F)}.$$

Indeed, for each $\lambda > 0$, we have $\alpha^0(\lambda, F) = 2(\lambda_F^0)^{-1}(\lambda) = \frac{1}{f(M_F)} f(c^0(\lambda, F) + \tfrac{1}{2}\lambda) = \frac{1}{f(M_F)} Q_F'(\lambda)$ (the relations $Q_F(\lambda) = F(c^0(\lambda, F) + \tfrac{1}{2}\lambda) - F(c^0(\lambda, F) - \tfrac{1}{2}\lambda)$ and $f(c^0(\lambda, F) + \tfrac{1}{2}\lambda) = f(c^0(\lambda, F) - \tfrac{1}{2}\lambda)$ give the last equality).

4. Tailweight and skewness with respect to the mode

We consider now some concepts of *tailweight* and qualitative aspect of *skewness with respect to mode*. Symmetry is a well-defined notion, whereas the concept of skewness is more complex; it depends in particular on the centre considered and on the weight used. Here the description is centered on the mode and the probability density function may be an appropriated measure of tailweight.

4.1. Tailweight orderings

Let us define the following tailweight ordering with respect to the mode; L_F denotes the length of S_F.

Definition 4.1.

$$F \leqslant_S^0 G \Leftrightarrow \lambda_G^0 \circ (\lambda_F^0)^{-1} \quad \text{convex on } [0, L_F].$$

In the symmetric case, we obtain: $F \leqslant_S^0 G \Leftrightarrow (\widetilde{G})^{-1} \circ \widetilde{F}$ convex on $(m_{\widetilde{F}}, +\infty) \cap S_{\widetilde{F}}$, i.e. $\widetilde{F} \leqslant_S \widetilde{G}$, where $\leqslant_S$ denotes van Zwet's (1964) tailweight ordering. The ordering $\leqslant_S^0$ allows to compare several usual distributions. We obtain straightforwardly:

Proposition 4.2.

- *uniform* $\leqslant_S^0$ *normal* $\leqslant_S^0$ *logistic* $\leqslant_S^0$ *double exponential.*
- *logistic* $\leqslant_S^0$ *Cauchy.*

So several usual distributions ordered for $\leqslant_S$ are also ordered by $\leqslant_S^0$. However, the implication $F \leqslant_S G \Rightarrow F \leqslant_S^0 G$ is not true, as illustrated by the following example.

Example 4.3. Let F be the symmetric triangular distribution function on $[-2, 2]$ and G the Cauchy distribution function; then we have $G \leqslant_S F$. However F is not $\leqslant_S^0$-comparable with G.

The implication $F \leqslant_S^0 G \Rightarrow F \leqslant_S G$ is an open question.

Remarks 4.4. Loh (1982, p. 27) proves that, if F is a symmetric unimodal and continuous distribution function, then F strongly unimodal implies $F \leqslant_S DE$, where DE is the double exponential law. We improve this result:

For any symmetric unimodal distribution function F,

$$F \text{ strongly unimodal} \Leftrightarrow F \leqslant_S^0 DE.$$

MacGillivray (1986) underlines that "van Zwet's skewness ordering has no reference to any measures of location and scale, and any weakening of the ordering in the sense of covering larger classes of distributions, involves reference to particular location and scale parameters." An analogous property occurs concerning tailweight. All weakened forms of $\leqslant_S$ are successively obtained, first by the choice of the position of the *shoulders* (points separating the central part of the distribution and the tails), and then by the choice of the scaling technique (Balanda and MacGillivray, 1990). In a mode-based approach, for $\alpha \in (0, \frac{1}{2}]$, $f_-^{-1}(\alpha)$ and $f_+^{-1}(\alpha)$ represent the shoulders of a probability density function f. A similar approach of tailweight orderings with respect to the mode is possible. Indeed, consider the following tailweight ordering with respect to the mode:

Definition 4.5. Let $\alpha \in (0, \frac{1}{2}]$;

$$F \prec_*^\alpha G \Leftrightarrow u \mapsto \frac{\lambda_G^0(u) - \lambda_G^0(\alpha)}{\lambda_F^0(u) - \lambda_F^0(\alpha)} \text{ nondecreasing on } (0, \tfrac{1}{2}].$$

Proposition 4.6.

$$F \leqslant_S^0 G \Leftrightarrow \forall \alpha \in (0, \tfrac{1}{2}], \qquad F \prec_*^\alpha G.$$

Proof. The result follows from the definition of $\leqslant_S^0$ and from the property that h is convex if and only if $x \mapsto \frac{h(x) - h(x_0)}{x - x_0}$ is nondecreasing, for each x_0. $\square$

Remarks 4.7. Let us note sG the symmetrized version of a distribution function G defined by Doksum (1975):

$$(sG)^{-1}(u) = \tfrac{1}{2}(G^{-1}(u) - \bar{G}^{-1}(u)), \quad u \in (0, 1),$$

where $\bar{G}(x) = 1 - G(-x)$. Let now f be a unimodal probability density function with mode 0, and let us consider the unimodal symmetric probability density function defined by:

$$f_S(x) = \begin{cases} s\tilde{F}(x) & \text{for } x \leqslant 0, \\ 1 - s\tilde{F}(x) & \text{for } x > 0. \end{cases}$$

f_S is called the *symmetrized version* (*with respect to the mode*) of f and let F_S be the associated distribution function. Further, let $\mathscr{D}$ denote the set of all symmetric unimodal distribution functions with mode 0; then $F_S \in \mathscr{D}$ and it minimizes $\tilde{\Delta}(F, G) = \sup_{u \in (0,1)} |\tilde{F}^{-1}(u) - \tilde{G}^{-1}(u)|$ for $G \in \mathscr{D}$ (Doksum, 1975).

Moreover, F_S is the unique element of $\mathscr{D}$ such that $\lambda_{F_S}^0 = \lambda_F^0$. The tailweight orderings $\prec$ with respect to the mode previously considered are expressed in terms of λ_F^0, so we have: $F \prec G \Leftrightarrow F_S \prec G_S$. Hence tailweight properties of F reduce to those of F_S. As for quantile-based quantities (Balanda and MacGillivray, 1990, p. 20), in the asymmetric case, it is difficult to separate differences of shape due either to skewness or to tailweight. Therefore the advantage of studying tailweight properties on symmetrized versions.

4.2. Skewness: qualitative aspect

We restrict our attention to the qualitative aspect of skewness with respect to the mode. However, it is likely to be possible to approach the comparative and quantitative aspects of skewness in the same way as for tailweight, using the existing results of skewness orderings and measures, and using the map $F \mapsto \tilde{F}$. Different notions of skewness to the right are to be found in the literature. We recall some of these concepts. The first and second are due to Doksum (1975) and the third one to van Zwet (1964).

Definition 4.8.

SK1 $\forall t > 0, \quad F(m_F(\frac{1}{2}) - t) + F(m_F(\frac{1}{2}) + t) \leqslant 1,$

SK2 m_F nonincreasing on $(0, \frac{1}{2}]$,

SK3 $F^{-1} \circ \tilde{F}$ convex (where $\tilde{F}(x) = 1 - F(-x)$).

Averous and Meste (1990) study these concepts in terms of properties of the midpoint of the median central interval. By analogy, we introduce definitions of *skewness to the right with respect to the mode*:

Definition 4.9.

SKM1* $\forall \lambda \geqslant 0, \quad c^0(\lambda, F) \geqslant c^0(0, F),$

SKM2* $c^0(\cdot, F)$ nondecreasing on $\mathbb{R}_+$,

SKM3* $c^0(\cdot, F)$ convex and nondecreasing on $\mathbb{R}_+$.

We have SKM3* $\Rightarrow$ SKM2* $\Rightarrow$ SKM1*.

Remarks 4.10. SKM1* can be rewritten in the following form:

SKM1 $\forall x > 0 \quad f(M_F - x) \leqslant f(M_F + x).$

Moreover, if we assume that F has a differentiable probability density function f, then definitions SKM2* and SKM3* reduce to:

SKM2 $\forall u \in (0, \frac{1}{2}] \quad f'(\tilde{f}_-^{-1}(u)) \geqslant -f'(\tilde{f}_+^{-1}(u)),$

SKM3 $u \mapsto -f'(\tilde{f}_-^{-1}(u))/f'(\tilde{f}_+^{-1}(u))$ nonincreasing on $(0, \frac{1}{2})$.

Remarks 4.11. SKM1 was used by Groeneveld and Meeden (1977) in order to obtain, for continuous probability density functions on $(0, +\infty)$, the "mean–median–mode inequality". Let us note SKM2$_s$ the SKM2 condition with strict inequality; Runnenburg (1978) considered strictly unimodal distributions, with differentiable probability density functions. For these distributions, he gave SKM2$_s$ on the one hand as sufficient condition for "mean–median–mode inequality" (Timerding's theorem, p.75), and on the other hand as sufficient condition for the positivity of odd central moments. (See also Bélisle, 1991.)

Now let us give some examples of skewed distributions:

Example 4.12. For each $0 < m < k$, the Beta(k, m) distribution is skewed to the left in SKM2 sense (Runnenburg, 1978). For each $\alpha > 1$, the Gamma(α, β) distribution is skewed to the right in the SKM3 sense.

Let us remark that skewness with respect to the median and with respect to the mode differ, even if median and mode coincide. This property is illustrated by the following example:

Example 4.13. Let f be the following probability density function:

$$f(x) = \begin{cases} 0 & \text{for } x \leqslant -\frac{3}{2} \text{ or } x \geqslant 1, \\ \frac{2}{5}x + \frac{3}{5} & \text{for } x \in [-\frac{3}{2}, -\frac{1}{4}], \\ 2x + 1 & \text{for } x \in [-\frac{1}{4}, 0], \\ -x + 1 & \text{for } x \in [0, 1]. \end{cases}$$

This probability density function is skewed with respect to the median in the SK1 sense, but presents no skewness in SKM1 sense.

However, we have the following property (van Zwet, 1979, p. 3):

$$\text{SKM2}_s \Rightarrow \text{SK1}.$$

Recall that in his paper, van Zwet gives SK1 as the sufficient condition for the "mean–median–mode inequality"; hence his theorem includes the result of Groeneveld and Meeden (1977) and of Runnenburg (1978).

Let us finally remark that SK3 and SKM2$_s$ are not comparable, neither one implies the other, as studied by Runnenburg (1978).

Acknowledgements

References

Averous, J. and M. Meste (1990), Location, skewness and tailweight in L_s-sense: a coherent approach, *Statistics* 21(1), 57–74.
Balanda, K.P. and H.L. MacGillivray (1988), Kurtosis: a critical review, *Amer. Statist.* 42, 111–119.
Balanda, K.P. and H.L. MacGillivray (1990), Kurtosis and spread, *Can. J. Statist.* 18(1), 17–30.
Bélisle, C. (1991), Odd central moments of unimodal distributions, *Statist. Probab. Lett.* 12, 97–107.
Bertin, E.M.J., W. Hengartner and R. Theodorescu (1981), Some characterizations of unimodal distribution functions, *Z. Wahrsch. Verw. Gebiete* 57, 327–338.
Bickel, P.J. and E.L. Lehmann (1979), IV spread, in: J. Jurečková, ed., *Contribution to Statistics*, Jaroslav Hájek Memorial Volume, pp. 33–40 (Riedel, Dordrecht).
Dharmadhikari, S. and K. Joag-Dev (1988), *Unimodality, Convexity and Applications* (Academic Press, New York).
Doksum, K.A. (1975), Measures of location and asymmetry, *Scand. J. Statist.* 2, 11–22.
Groeneveld, R.A. and G. Meeden (1977), The mode, median and mean inequality, *Amer. Statist.* 31, 120–121.
Groeneveld, R.A. and G. Meeden (1984), Measuring skewness and kurtosis, *The Statistician* 33, 391–399.
Loh, W.Y. (1982), Tail-orderings on symmetric distributions with statistical applications, Thesis, Univ. California, Berkeley.
MacGillivray, H.L. (1986), Skewness and asymmetry: measures and orderings, *Ann. Statist.* 14(3), 994–1011.
Runnenburg, J.Th. (1978), Mean, median, mode, *Statist. Neerlandica* 32, 73–79.
van Zwet, W.R. (1964), *Convex Transformations of Random Variables*, Math. Centre Tracts 7 (Mathematisch Centrum, Amsterdam).
van Zwet, W.R. (1979), Mean, median, mode II, *Statist. Neerlandica* 33, 1–5.